1. REPORT DATE **APR 2004**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**	
4. TITLE AND SUBTITLE **How Compensation in Test and Evaluation Affects Aircraft Acquisition**		5a. CONTRACT NUMBER	
		5b. GRANT NUMBER	
		5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)		5d. PROJECT NUMBER	
		5e. TASK NUMBER	
		5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defense Acquisition University Alumni Association 2550 Huntington Ave, Suite 202 Alexandria, VA 22303**		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)	
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**			
13. SUPPLEMENTARY NOTES			
14. ABSTRACT			
15. SUBJECT TERMS			

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT **UU**	18. NUMBER OF PAGES **14**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

HOW COMPENSATION IN TEST AND EVALUATION AFFECTS AIRCRAFT ACQUISITION

Lt Col Lionel D. Alford, USAF

Systems developers and testers have assumed that human compensation is measurable, or at least that a cognizant and trained tester is able to identify and detect compensation. More than one study conducted at the Wright-Patterson LAMARS facility indicates that this is not necessarily true. Test pilots were able to compensate sufficiently to fly and meet defined performance standards on intentionally crippled aircraft flight control designs. These flight control systems were designed to trigger pilot induced oscillations, but in most cases, test pilots could compensate sufficiently to prevent pilot induced oscillations and to control the simulated aircraft. Test pilot compensation hides critical handling qualities cliffs that can lead to loss of an aircraft when encountered by less skilled pilots. This observation has vast ramifications for test, evaluation, and development of all human interface systems.

After solving the problems of propulsion and lift, the control of an aircraft was the third and possibly greatest challenge the Wright Brothers faced in conquering the air (Figure 1). After all, the Wright Brothers really took eight years following their historic first flight to determine the problem of stall in a turn and how to correct it (Culick, 2001). They knew how to control the aircraft, but the controls were insufficient. The aircraft design, power, and stability were all factors in solving that problem.

Because aircraft design, power, and stability all affect and are affected by the aircraft control system, the key factor in the development of control systems for aircraft is to design them to optimize the aircraft performance while providing *carefree* handling qualities to the pilot. For example, the best fuel economy is achieved when the center of gravity is behind the center of lift. The center of lift is the neutral point of the aircraft and having the center of gravity at or behind the center of lift creates a condition where

---

DISCLAIMER

The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government.
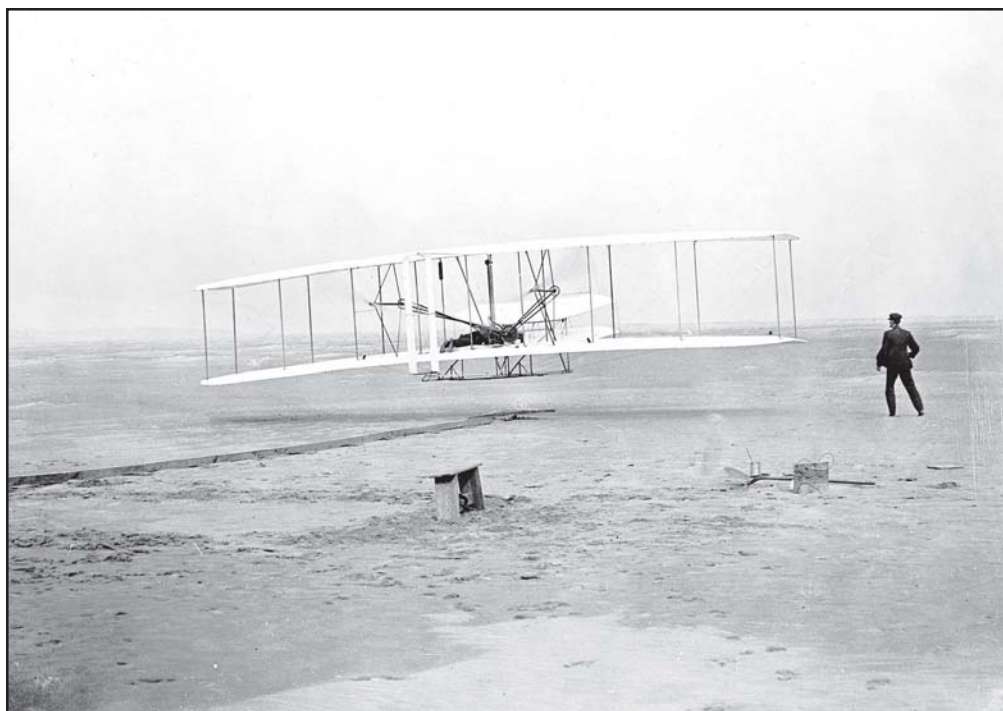

Figure 1. The Wright brother's first flight was still eight years away from a fully stable/controllable aircraft.

the aircraft is neutrally stable or unstable — both bad for handling qualities.

An aircraft can be designed with an electronic flight control system that gives the pilot a positive, stable, handling-qualities feel while the aircraft is unstable (Baer & Landy, 1987). This design results in increased fuel economy, a characteristic of the Airbus 320+ and the Boeing 777. Future airliners will certainly capitalize and expand on this capability. Another example of unstable aircraft design to achieve aircraft requirements is found in modern fighter-type aircraft. For radar stealth, maneuvering performance, and mission optimization, among other reasons, the exterior of military aircraft are designed in such a way that, without electronic flight controls, they would be unflyable (Rushby, 1993). The F-16, F-18, F-22, F-117, and B-2 are all examples of this type of design (Rushby, 1993).

The problem of aircraft handling exists because aircraft controls are counter-intuitive. Water and land-borne transportation turn using a device like a rudder to modify the velocity vector in the horizontal plane. This is an intuitive response that is easy to master. To turn an aircraft requires a roll in the horizontal plane coupled with a pitch rotation to counter loss of lift in the vertical plane and an increase in thrust to balance the increase in drag. The pitch rotation, and not the roll, turns the aircraft. A coordinated turn further requires a corrective yaw rotation in the horizontal plane to counter the slip induced by the original roll.

Aircraft motion is also characterized as a mass-spring-damper and therefore is a

system that responds in a manner significantly different than ground-based controls. The problem of aircraft control in thrust and pitch is further complicated by the power curve response in the region of reverse command where the pitch control largely directs airspeed and the thrust directs pitch — again an unintuitive response.

People can generally be trained to adequately control unintuitive systems such as aircraft. It should be self-evident that both ground and flight control systems represent natural phenomena that are within easy grasp of human beings. However, higher-order systems, those greater than second order, are not generally found in nature and may not be predictable by human beings (National Research Council, 1997; Rushby, 1993).

Modern digital flight control systems use approximations of mathematical equations that result in very high order systems to replicate the natural response of a non-electronically controlled aircraft. These systems of high-order approximations generally do a great job of reproducing the handling qualities of the perfect aircraft; however, they result in a system that potentially is unpredictable to the operator and they introduce unpredictable response in the overall aircraft system (National Research Council, 1997).

The problem of unpredictability of a flight control system is characterized by handling quality cliffs and pilot induced oscillation (PIO). A handling quality cliff is an unknown and untested area in a flight control envelope where it is possible for the pilot to unexpectedly lose control of the aircraft. A PIO is a situation where aircraft response lags the pilot's input to the controls. The pilot unconsciously increases control input such that each input magnifies the aircraft response until loss of control or the aircraft comes apart. PIO is not unique to digital flight control systems, but unforeseen PIOs are. These problems are best characterized by the distinctive mishaps they have spawned.

On October 26, 1977, the prototype Space Shuttle was launched from its 747 carrier aircraft. The pilots, Fred Haise and Gordon Fullerton, attempted a spot landing on the concrete main runway at Edwards Air Force Base. The shuttle had an electronic triply redundant digital fly-by-wire flight control system. The expected performance of the aircraft did not match the actual performance and Haise found himself too fast on the approach. His overcompensation resulted in a PIO in roll and pitch. In spite of this, he landed the shuttle safely (STS Approach and Landing Test, 1977; STS Space Shuttle, 1977).

"It should be self-evident that both ground and flight control systems represent natural phenomena that are within easy grasp of human beings."

NASA engineers found a 270-millisecond time delay in the flight control system that they corrected with a filter (STS Space Shuttle, 1977). In spite of this change, the pilot astronauts know the shuttle cannot be flown like a fighter.

The flight control system of the shuttle was based on and is similar to the F-16. The F-16 had and has a known 270-millisecond time delay in the pitch axis. If the aircraft is mishandled, this delay will result in a PIO, and PIO has been the focus of numerous mishap investigations (Rushby, 1993). The result is that pilots fly an F-16 approach and

landing like they would a heavy aircraft and not like a fighter.

Usually, critical PIO problems are only identified as a problem when they cause an aircraft mishap. On February 2, 1987 during its seventh flight, the first prototype Swedish Saab JAS39 Gripen crashed on landing. The Gripen had a triplex redundant fly-by-wire digital control system backed up by a triplex redundant analogue fly-by-wire control system. The first test pilot remarked that the flight control system was too sensitive and displayed problems with lateral and pitch oscillations. The pilot flying during the mishap had never flown the Gripen, and gusty wind conditions likely exacerbated the problems with lateral and pitch oscillations. The pilot encountered increasing PIOs characterized by dynamic pitch instability during approach. These control problems resulted in the aircraft striking a wing on landing and the destruction of the aircraft (Aviation Week, 1989; Flight International, 1989; Nutley, 1989; Pellebergs, 1991).

The Gripen program went through a very intensive flight and ground test program to fix the problems caused by the flight control system, and the aircraft continued development. Everything appeared fine until on August 8, 1993, during a normal maneuver, a pilot flying the Gripen in an airshow fully saturated the flight controls and entered an unrecoverable PIO. The manufacturer and the customer knew that large stick movements could saturate the flight control system, but the pilot was unaware of this aircraft characteristic. The aircraft was destroyed (Swedish Accident Investigation, 1993).

The Gripen is not the only aircraft that has experienced interaction of the pilot

Figure 2. The latest F-22 loss of control incident is possibly a digital flight control cliff.

and the flight control system that resulted in the loss of a prototype aircraft. On April 25, 1992, the YF-22 (the production F-22 is shown in Figure 2) also crashed during landing due to PIO caused by the fly-by-wire system (Rushby, 1993).

Although PIO is a known problem of non-fly-by-wire flight control systems, time delays, handling qualities cliffs, unpredictable flight characteristics, saturated control systems, the attendant PIOs, and loss of control are characteristics of fly-by-wire flight control systems. One final example will illuminate this problem of PIOs as it relates to testing. Many of these control problems manifest themselves in the roll axis. In heavy aircraft, the problems result in a faster roll rate than normally expected (Norton, 1994). This unpredictability, combined with transport delays cause PIOs. The C-17 program encountered this problem late in its test program. The aircraft had a known quick roll rate, but test pilots who had been flying the aircraft for a while did not consider it a problem. New test pilots generally complained about the roll rate and its attendant PIO during landing approach, but they quickly learned to compensate. When a test pilot new to the C-17 recognized the problem and complained officially, the program blamed the pilot and continued with the control system unchanged. In Operational Test & Evaluation (OT&E), the operational pilots reported the problem as an aircraft deficiency and that is when it was finally fixed.

The difficulty in the C-17 program wasn't simply that a handling qualities problem existed in the roll axis. The problem was that so many trained test pilots, military and civilian, had flown an aircraft with an obvious deficiency and found it acceptable without changes. The factor of aptitude that allows trained test pilots to compensate for evident deficiencies in flight control system is the problem this paper directly addresses.

THE EXPERIMENT

The above situations demonstrate that PIOs in fly-by-wire aircraft are neither uncommon nor insignificant, and they are more likely to be experienced by non-test pilot aviators. At best, they represent a nuisance and, at worst, a potentially catastrophic air event. Because of these problems, their criticality to flight and the danger of not finding them during developmental testing, we are studying PIOs, when possible, through simulations.

In December 1998, I participated as a subject in the HAVE LIMITS SIM and HAVE PIO SIM, PIO study conducted by the Air Force Research Laboratory Air Vehicle Directorate (AFRL/VA) that used the Large Amplitude Multi-Mode Aerospace Research Simulator (LAMARS). LAMARS is a 20-foot diameter sphere on the end of a 30-foot beam that comprises a 5-Degree of Freedom Simulator (Figure 3). The simulator includes a McFadden Feel System, wrap-around

> *"The factor of aptitude that allows trained test pilots to compensate for evident deficiencies in flight control system is the problem this paper directly addresses."*
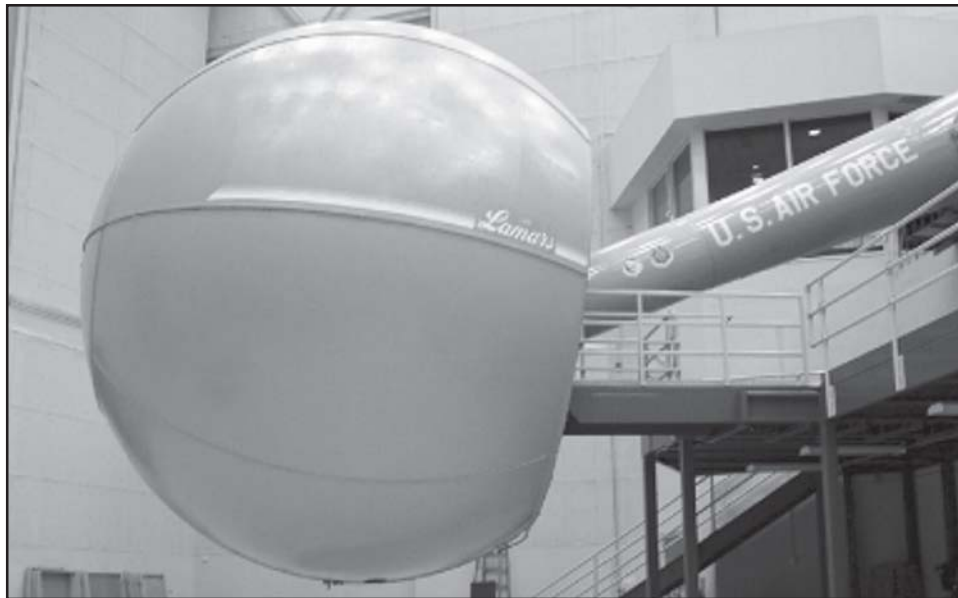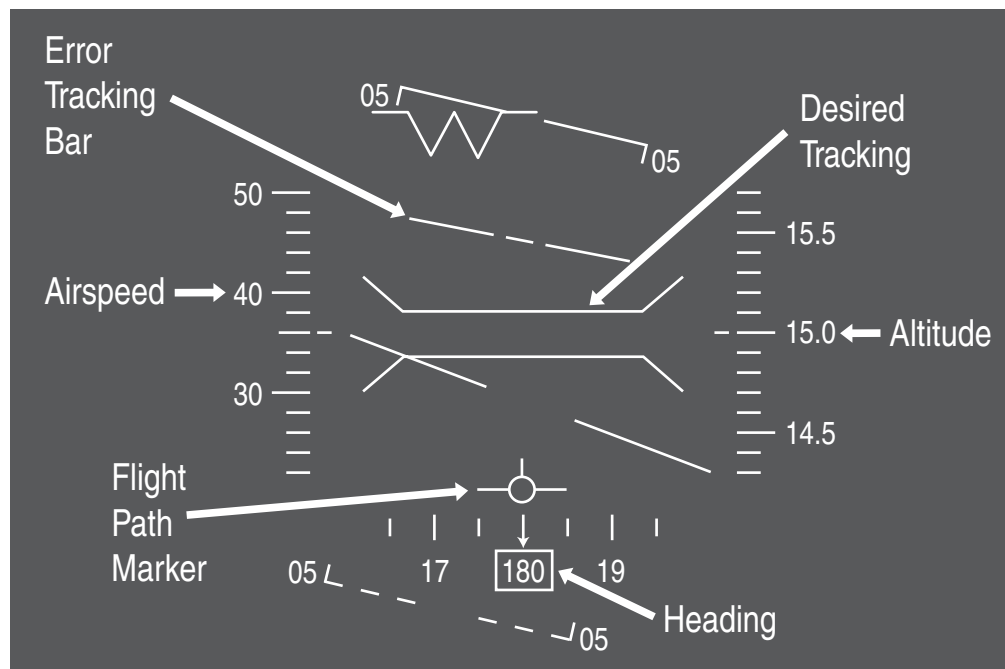
Figure 3. The LAMARS simulator

Figure 4. The heads-up display during the LAMARS test

visuals, and a Heads-Up Display (HUD) (Figure 4). The LAMARS system is capable of up to –2 to +3 g vertical and -3 to +3 g horizontal acceleration.

The purpose of the experiment was to gather data on aircraft handling qualities models (good, bad, and ugly) to correlate in-flight variable simulation data from the original HAVE LIMITS and HAVE PIO open air programs with simulation data. The experiment featured three pitch and roll capture tasks with increasing levels of workload and a landing task. The aircraft handling qualities models varied widely based on 18 variations tested previously in the CALSPAN/Veridian NT-33 variable stability aircraft. The study itself resulted in good data but simulation results could not be correlated with open-air flight test (Stadler, 1999). The observations I made came out of a deviation from the test and were not included directly in the test or research.

I had the opportunity to be the first of three test pilots who participated in the program. When I arrived to make the first runs, the LAMARS was not ready for motion. I flew a full set of the simulations as training without simulator motion. A few days later, I flew the remaining training and data runs with LAMARS accelerative motion. My observations come from the unique perspective of being able to see how the simulations flew with and without accelerative motion.

When the motion was off, some of the configurations were impossible to fly. Many of the flight control designs were divergent and resulted in complete loss of control. The simulator was relatively easy to PIO and many of the runs resulted in a departure from controlled flight. In most cases, the pilot could not respond quickly enough to go open loop when a PIO was immanent. Further, as reason might indicate, the higher the workload of the task the easier it was to depart the system. This was not true during the simulation runs with accelerative motion.

When accelerative motion was on, the tasks became easier to fly as the workload increased. Pilot compensation and learning occurred at a rate not possible without motion. Due to the natural feel in the acceleration, it was increasingly easy with increasing workload to maintain control of the aircraft. The pilot had to force himself to allow PIO conditions to continue. It was very easy to reduce workload slightly and allow the system to dampen out instead of pulling aggressively to the point that would have departed the system. This was very obvious with negative G during pushovers. Although, the simulations felt like they were often on the ragged edge of departure, it was possible to prevent a PIO and a departure. The feedback I received was that the departure rate overall during the study was lower than expected, "the bad was not as bad" as seen in the actual aircraft (Stadler, 1999). Additionally, the researchers observed that pilot anticipation of the PIOs may have skewed the data.

It would be easy at this point to conclude that the study itself did not provide much useful data, but I think this

> **"When accelerative motion was on, the tasks became easier to fly as the workload increased."**

program highlighted a very critical area that has not been considered much in flight testing — test pilots may not be able to gauge how much they are compensating.

RESULTS

The observation that *test pilots may not be able to gauge how much they are compensating* is not as obvious as it seems. We expect average pilot subjects to not know they are compensating — we assume they are compensating and we assume we can measure that compensation through workload. For test pilots the situation is different. A key skill in which we train test pilots is to observe and know when they are compensating. This is the proverbial test pilot *handshake*. If test pilots cannot gauge their compensation, then there is little hope of solving the critical problems that face us in digital flight control systems. Indeed, based on this observation, we may have to look for a different way of designing and testing not just flight control systems but all types of human-machine interfaces.

As long as the aircraft is predictable, and predictability increases dramatically with natural accelerative motion, the pilot can apply reflexive filtering that normally prevents PIO and departure. As my experience in the simulator with and without accelerative motion demonstrates, without accelerative motion, the system is not as predictable as with accelerative motion. Without G force, the system is less predictable. The aircraft's acceleration makes

"Acquiring firms consider both distressed and highly profitable firms as potential acquisition candidates."

possible *heroic response* to bad flight control systems.

I assert *heroic response* with intended experimental precision. Heroic response is exactly what any pilot accomplishes when faced with a poor flight control system design. Experienced pilots unconsciously feel the natural/predictable modes of an aircraft and successfully compensate for poor handling qualities. Most pilots do not realize the degree of compensation used to counteract normal aircraft handling qualities.

In an aircraft development program, as an aircraft flight control system improves, the test pilot's compensation improves and without a significant event, such as a recalibration of the pilot's compensation experience, the compensation will continue to improve. As with the C-17 example, and much of my flight test experience shows, test pilots, like all pilots, will at some point no longer be able to gauge their compensation and then they will not be of much use to the test program. Without training or comparisons, it may be impossible for pilots to gauge the degree of compensation, especially with long-term programs and programs where handling qualities have improved gradually over time.

This observation is true of flight control systems as well as any other control system in an aircraft. I further suspect that this observation concerning compensation and test pilots is true of tests in all other complex systems.

In the case of unnatural or unpredictable modes of digital flight control aircraft, these modes can only be learned through experience — if undiscovered and uncorrected, these handling qualities *cliffs* will result in loss of aircraft. These characteristics of pilot compensation make digital flight control aircraft more difficult to sufficiently test.

Pilots may learn to compensate to the degree that they unconsciously filter even unnatural and unpredictable modes. However, if the mode is not experienced, is unpredictable, or is not discovered and corrected during testing, some operational pilot will eventually encounter a handlings quality cliff, and recovery may not be possible.

RECOMMENDATIONS

1. More research needs to be accomplished on measuring pilot compensation — and workload may not be a good measure. Workload measurement has been the *Holy Grail* of human factors testing. To date we do not have a quantitative measure of workload and this makes human factors testing subjective and difficult. Quantitative workload measurement is a needed and necessary tool for human-machine interface development, but there is a piece of the puzzle that is still missing in workload measurement — how do we quantify compensation?

2. Test pilots require hands-on training to understand the level of compensation possible during test programs. The LAMARS facility with its PIO models provides an excellent means of training. This training should be required in every Test Pilot School curriculum and taught as continuing Test Pilot education. Prior to the test of digital flight control system aircraft, the pilots on the program should attend some level of refresher orientation. The training should allow the comparison between *seat-of-the-pants* accelerative motion and no-motion to drive home the point

that too much exposure to bad flight control models skews the pilot's perspective, and a pilot can become too comfortable with a poor flying system.

3. Test pilots need to constantly recalibrate their awareness of aircraft handling quality differences and compensation. The best way to achieve this is through multiple qualifications and qualification flights in different aircraft. All programs could benefit from this regimen. Test pilots who don't fly multiple aircraft and who cannot compare different designs and systems lose the ability to identify their level of compensation. The best method to keeping this critical skill sharp is to experience known deficient designs and poor handling aircraft. Test pilot schools and test centers should ensure a large number of poor aircraft and historical aircraft are available for test pilot qualification. The services should address this problem by allowing test pilot access to more systems.

4. An obvious but often overlooked recommendation in the analysis of fly-by-wire systems is that developers should attempt to design predictable flight models that don't just mimic natural aircraft response but truly match it.

CONCLUSION

We may have underestimated the role of compensation in testing and we need to determine ways to measure compensation. Pilots can learn to adequately fly poor aircraft with intentionally poor handling qualities. They appear to be able to unconsciously filter certain characteristics in the handling

qualities envelope of the aircraft. Unfortunately, systems developers and testers have always assumed that human compensation is measurable, or at least that a cognizant and trained tester is able to identify and detect compensation. The HAVE LIMITS SIM and HAVE PIO SIM studies conducted at the Wright-Patterson LAMARS facility indicate that this is not necessarily true. Test pilots were able to compensate sufficiently to fly and meet defined performance standards on intentionally crippled aircraft flight control designs. This creates critical questions for the testing of future human interface system. To help solve these problems:

- More research needs to be accomplished on measuring pilot compensation.

- Test pilots require hands-on training to understand the level of compensation possible during test programs.

- Test pilots need to constantly recalibrate their awareness of aircraft handling quality differences and compensation.

- Developers should attempt to design predictable flight models that don't just mimic natural aircraft response but truly match it.

Digital control systems create unique problems for engineering design and flight. In the case of aircraft, the best design approach may be to develop predictable flight models that directly match or simply augment natural aircraft response instead of using complex digital equations that imitate assumed aircraft response. Until that point is reached and because of the difficulties involved in designing human interfaces and the human control of complex systems, we must find quantitative ways to measure compensation and we must control experiments to address compensation issues.

Lt Col Lionel D. Alford, Jr., USAF, is the chief of system safety for the Global Reach System Program Office at the Aeronautical System Center at Wright-Patterson AFB, Ohio. He is an Air Force experimental test pilot with over 4900 hours in more than 60 different kinds of aircraft and is a member of the Society of Experimental Test Pilots. Alford holds an Airline Transport Pilot license and was the chief test pilot for a number of Air Force acquisitions. He is a graduate of the Defense Systems Management College Advanced Program Management Course, Air War College, and Air Command and Staff College. He is a doctoral candidate in aeronautical engineering at the University of Dayton and holds a M.S. in mechanical engineering from Boston University and a B.S. in chemistry from Pacific Lutheran University.

(E-mail address: Lionel.Alford@wpafb.af.mil)

BIBLIOGRAPHY

A330 accident preliminary report summary. (1994). *Air et Cosmos*, *1482.*

A330 accident preliminary-Preliminary findings from the telemetry data. (1994). *Air et Cosmos*, *1480.*

Angner, J., Jensen, C., & Seidl, M. (1996). JAS 39 Gripen EFCS: How to deal with rate limiting. *SETP 40 Symposium Proceedings*.

Beatson, J. (1989, April 2). Is America ready to "Fly by Wire"? *The Washington Post*, p. C3.

Gerner, A., & Manning, M. (1991). Frequency response analysis and modi-
fied maximum likelihood estimation of a T-38A. Edwards AFB, CA: USAFTPS.

Moss, C., & Shifflett, S. (1997). B-2A Residual Pitch Oscillation (RPO) investigation. *SETP 41 Symposium Proceedings,* Edwards AFB, CA: 420 Flight Test Squadron.

REFERENCES

Baer-Riedhart, J., & Landy, R. (1987). Highly integrated digital electronic control — digital flight control, aircraft model identification, and adaptive engine control. NASA Technical Memorandum 86793. Edwards AFB, CA: NASA Ames Research Center, Dryden Flight Research Facility.

Staff. (1989, February 27). Saab blames Gripen crash on software. (1989, February 27). *Aviation Week & Space Technology*, 130 (9), 31.

Culick, F. E. C. (2001). Wright brothers: First aeronautical engineers and test pilots. California Institute of Technology: *SETP 45 Symposium Proceedings*. Retrieved December 30, 2003 from http://www.wrightflyer.org/Papers/SETP01_Culick.pdf

Department of Defense, Defense Management Review. (1989, July). Defense management: Report to the President. Secretary of Defense, Richard Cheney. Retrieved February 29, 2004 frorm http://www.defenselink.mil/nii/bpr/bpred/289a.htm

National Research Council. (1997). Aviation safety and pilot control: Understanding and preventing unfavorable pilot-vehicle interactions. National Academy Press. Washington, D.C.

Norton, W. J., & Melroy, P. A. (1994). C-17A airloads flight tests. *SETP 38 Symposium Proceedings,* Edwards AFB, CA: 412 Test Wing, Air Force Flight Test Center.

Nutley, M. (April 3, 1989). "Swedish wind cuts fly-by-wires." Datalink (a British paper for system software). Retrieved March 3, 2004 from http://catless.ncl.ac.uk/Risks/8.49.html

Pellebergs, P., & Lindholm, A. (1991). JAS 39 Gripen flight test report. *SETP 35 Symposium Proceedings.* SAAB-Scania AB, Aircraft Division.

Rushby, J. (1993, November). Anomalies in digital flight control systems (DFCS): SRI international computer science laboratory. Report SRI-CSL-93-07.

Stadler, B. (1999). Recommendations to improve future pilot-induced oscillation simulations. Wright-Patterson AFB, OH: AFRL/VACD.

STS Approach and Landing Test (ALT). (1977, February). First Shuttle/747 captive flight. Retrieved December 29, 2003 from http://www.dfrc.nasa.gov/gallery/photo/ALT/HTML/ECN-6887.html

STS Space Shuttle. (1977). STS approach and landing test (ALT): Flight 5 — pilot-induced oscillation (PIO) on landing: Movie Collection. Retrieved December 29, 2003 from http://www.dfrc.nasa.gov/gallery/movie/STS/HTML/EM-0043-02.html

Swedish Accident Investigation. (1993). The Swedish accident investigation boards report summary. JAS 39 Gripen crash in Stockholm 1993 Aug 08 report summary. Retrieved December 30, 2003 from http://www.canit. se/~griffon/aviation/text/gripcras.htm

Gripen crash caused by flight control softwarefighter crash. (1989, March 25). *Flight International*. Retrieved March 3, 2004 from http://catless. ncl.ac.uk/Risks/8.49.html